**IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF VIRGINIA
ALEXANDRIA DIVISION**

DIALECT, LLC,	)	
	)	
Plaintiff,	)	
	)	
v.	)	Civil Action No. 1:23-cv-581-TSE-JFA
	)	
AMAZON.COM, INC., *et al.*,	)	**JURY TRIAL DEMANDED**
	)	
Defendants.	)	
	)	

**OPENING BRIEF IN SUPPORT OF AMAZON'S MOTION
TO DISMISS PLAINTIFF'S AMENDED COMPLAINT
UNDER FEDERAL RULE OF CIVIL PROCEDURE 12(b)(6)**

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page(s)

TABLE OF AUTHORITIES
(CONTINUED)

Page(s)

I.   INTRODUCTION

Since the earliest days of the computer era, generations of scientists strived to develop a system that can understand and respond to the spoken word, many believing the goal was unattainable. Only recently—through intense research investment and the advent of cloud computing—leading technology companies have achieved what some thought was impossible, creating systems (*e.g.*, Amazon's Alexa virtual voice assistant) that can understand and respond to a wide range of human commands.

Dialect's patents in this case claim the result of that hard work of others while contributing no technology to achieve it. Each patent is directed to interpreting or responding to a spoken request. U.S. Patent Nos. 9,031,845 (the "'845 patent"), 7,693,720 (the "'720 patent"), 8,015,006 (the "'006 patent"), and 9,495,957 (the "'957 patent") (collectively, the "Context Patents") claim understanding and responding to a spoken request using its context. U.S. Patent Nos. 9,263,039 (the "'039 patent") and 8,195,468 (the "'468 patent") (collectively, the "Multimodal Patents") claim responding to a request that has spoken and non-spoken portions.[1] None claims or describes a specific technological solution for achieving any of these goals. But all disclose that they attempt to use computers to mimic the way that humans receive and respond to information in everyday conversation. In other words, they attempt to claim human conversations on computers but do not disclose how to do so. The patents thus attempt to monopolize decades of groundbreaking research by true innovators and the solutions yet to be devised by them. The patents are invalid under 35 U.S.C. § 101 and none of Dialect's allegations in the amended complaint, filed to overcome Amazon's first motion to dismiss on the same ground, changes this conclusion.

---

[1]   The patents at issue in this motion are attached to Dialect's complaint as Dkt. 1-2 ('720 patent), Dkt. 1-3 ('006 patent), Dkt. 1-5 ('468 patent), Dkt. 1-6 ('845 patent), Dkt. 1-7 ('039 patent), and Dkt. 1-8 ('957 patent).

Amazon thus renews its request that the Court dismiss Dialect's claims under these six patents with prejudice under Rule 12(b)(6) for failure to allege infringement of a patentable claim.

II.   FACTUAL BACKGROUND

Plaintiff Dialect, LLC ("Dialect") asserts seven patents against Amazon.com, Inc. and Amazon Web Services, Inc. (collectively, "Amazon"). (Compl. ¶ 1.) Six patents are at issue in this motion.[2] All six patents were filed by third-party VoiceBox Technologies, share much of their specifications, and list many of the same inventors.[3]

Dialect broadly accuses "Amazon's Alexa virtual assistant technology and offerings" spanning five categories of products and services: (1) the Amazon Echo smart speakers that support voice interaction with Alexa; (2) the cloud-based Alexa Voice Services that the smart speakers use to interact with Alexa; (3) the Alexa Auto software development kit (SDK) and Alexa Mobile Accessory Kit used to integrate Alexa features into vehicle infotainment systems and Bluetooth headphones and speakers, respectively; (4) the Alexa, Music, Shopping Apps, and the Amazon.com website that purportedly allow interaction with Alexa; and (5) "any other device, app, or instrumentality that includes, provides access to, or works with Alexa." (Compl. ¶¶ 3 n.1, 52, 57, 166.)

Amazon filed its initial motion to dismiss under Section 101 on July 17, 2023. (Dkt. 33.) Dialect did not file an opposition to Amazon's motion, instead opting to file an amended complaint on July 31, 2023. (Dkt. 35 ("Am. Compl.").) Dialect's amended complaint does not assert any

---

[2]   The seventh asserted patent, U.S. Patent No. 8,140,327 (the "'327 patent"), is directed to filtering noise from a spoken request and is not at issue in this motion. Nothing in this motion is an admission or implication that the '327 patent is directed to patent-eligible subject matter; it is not.
[3]   For example, Robert Kennewick is listed as an inventor on the '845, '720, '006, '468, and '039 patents. Richard Kennewick is listed as an inventor on the '845, '720, '006, '468, and '957 patents. David Locke and other individuals are listed as common inventors on the '845, '720, and '006 patents. Chris Weider is listed as an inventor of the '468 and '957 patents. Philippe Di Cristo is listed as an inventor of the '468 and '039 patents.

new patents or claims, but includes: new allegations regarding the purported business activities of a third party, VoiceBox; new allegations regarding business meetings between VoiceBox and Amazon; modified descriptions of the patents and their claims; and quotes of patent examiners' reasons for allowing the claims during prosecution. (*See* Am. Compl. ¶¶ 20-38, 42-48, 51-57, 60-64, 67-72, 75-82, 85-90, 93-96.)

A.   The Context Patents

The Context Patents (the '845, '720, '006, and '957 patents) fall into two categories. The first category includes the '845, '720, and '006 patents, each identically titled "Systems and Methods for Processing Natural Language Speech Utterances with Context-Specific Domain Agents." The '845 patent is a continuation of the '720 patent and both patents have an identical specification. The '006 patent shares much of its specification with them.[4] All three patents are generally directed to recognizing the words of a spoken request, interpreting the meaning of those words by determining the appropriate context, and responding to the request. (*See* Am. Compl. ¶¶ 45, 54, 79.) The '845 and '720 patents recite accomplishing this goal in a vehicle.

The second category, the '957 patent, likewise shares much of its specification with the '845, '720, and '006 patents. Like those patents, the '957 patent is also generally directed to receiving a spoken request and identifying the context of the recognized words but requires determining how those words compare to contexts associated with prior spoken requests using "rank scores." (*See* Am. Compl. ¶ 95.)

In its complaint, Dialect asserts claim 1 of the '845 patent, claim 1 of the '720 patent, claim 5 of the '006 patent, and claim 1 of the '957 patent. (Am. Compl. ¶¶ 104, 163, 273, 331.) Neither claim recites a technological solution. All are purely functional and reference only generic

---

[4]   Unless otherwise noted, all cited disclosures in the '845 patent also appear in the '720, '006, and '957 patents.

hardware components and black-box software modules. For example, claim 1 of the '845 patent recites a generic "mobile system" with generic "physical processors at a [generic] vehicle" that are somehow "programmed" to perform predictable functions that one would use in a human conversation: receive a spoken utterance; perform speech recognition on the utterance; parse and interpret the utterance; determine a domain and context; formulate a command based on the domain and context; determine whether the command is to be executed in the vehicle or not; and, based on that determination, either execute the command in the vehicle or ask an outside device to execute it:

> 1. A mobile system for processing natural language utterances, comprising:
>
> one or more physical processors at a vehicle that are programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:
>
> > *receive a natural language utterance* associated with a user;
> >
> > *perform speech recognition* on the natural language utterance;
> >
> > *parse and interpret* the speech recognized natural language utterance;
> >
> > *determine a domain and a context* that are associated with the parsed and interpreted natural language utterance;
> >
> > *formulate a command or query* based on the domain and the context;
> >
> > *determine whether the command or query is to be executed on-board or off-board* the vehicle;
> >
> > *execute the command or query at the vehicle* in response to a determination that the command or query is to be executed on-board the vehicle; and
> >
> > *invoke a device that communicates wirelessly over a wide area network to process the command or query* such that the command or query is executed off-board the vehicle in response to a determination that the command or query is to be executed off-board the vehicle.

The claim, while lengthy, recites only the aspirational goal of using generic processors "programmed"—in an unspecified way—to receive and recognize speech, determine its context, and execute the command.

Claim 1 of the '720 patent is similar. It recites a generic "mobile system" in a generic

vehicle that performs many of the same expected speech recognition, interpretation, and response

functions on generically-labeled black-box components: (1) a "speech unit" to receive a spoken

request; and (2) a "natural language speech processing system" with (a) a "speech recognition en-

gine" to recognize words based on the current or past dialogs; (b) a "parser" that interprets the

words using context and transforms them into a question; and (c) an "agent architecture" that gen-

erates a response to the question. ('720 patent at claim 1.) The '720 patent also recites functional

software elements called "domain agents" that are purportedly used somehow to help recognize

words and interpret the spoken request:

> 1. A mobile system responsive to a user generated natural language speech utterance, comprising:
>
> a *speech unit* connected to a computer device on a vehicle, wherein the speech unit *receives a natural language speech utterance* from a user and *converts the received natural language speech utterance into an electronic signal*; and
>
> a *natural language speech processing system* connected to the computer device on the vehicle, wherein the natural language speech processing system *receives, processes, and responds to the electronic signal* using data received from a plurality of domain agents, wherein the natural language speech processing system includes:
>
>> a *speech recognition engine that recognizes at least one of words or phrases* from the electronic signal using at least the data received from the plurality of domain agents, wherein the data used by the speech recognition engine includes a plurality of dictionary and phrase entries that are dynamically updated based on at least a history of a current dialog and one or more prior dialogs associated with the user;
>>
>> a *parser that interprets the recognized words or phrases*, wherein the parser uses at least the data received from the plurality of domain agents to interpret the recognized words or phrases, wherein the parser interprets the recognized words or phrases by:
>>
>>> *determining a context* for the natural language speech utterance;
>>>
>>> *selecting at least one of the plurality of domain agents* based on the determined context; and
>>>
>>> *transforming the recognized words or phrases into at least one of a question or a command*, wherein the at least one question or command is formulated in a grammar that the selected domain agent uses to process the formulated question or command; and
>>
>> an *agent architecture* that communicatively couples services of each of an agent manager, a system agent, the plurality of domain agents, and an agent library that

includes one or more utilities that can be used by the system agent and the plurality of domain agents, wherein the selected *domain agent uses the communicatively coupled services to create a response to the formulated question or command* and format the response for presentation to the user.

Putting aside the verbose nature of claim 1 of the '720 patent, none of the "units," "modules," "architectures," or "agents" refers to a concrete hardware or software implementation, and as discussed below, none is new. The claim is purely functional.

Like the '845 and '720 patent claims, claim 5 of the '006 patent recites a method of receiving, recognizing, and parsing a spoken request; formulating the request based on its context; and responding to the request. ('006 patent at claim 5.) It adds that the request is formulated according to a "grammar" used by a generic domain agent; undefined "prior probabilities or fuzzy possibilities" are used, somehow, to make inferences about what is being requested; and the inferences are transformed into unspecified software "tokens" to be used—in an unspecified way—by the domain agent to generate a response:

> 5. A method for processing natural language speech utterances with context-specific domain agents, comprising:
>
> *receiving*, at a speech unit coupled to a processing device, a natural language speech utterance that contains a request;
>
> *recognizing*, at a speech recognition engine coupled to the processing device, one or more words or phrases contained in the utterance using information in one or more dictionary and phrase tables;
>
> *parsing*, at a parser coupled to the processing device, information relating to the utterance to determine a meaning associated with the utterance and a context associated with the request contained in the utterance, wherein the parsed information includes the one or more recognized words or phrases;
>
> *formulating*, at the parser, the request contained in the utterance in accordance with a grammar used by a domain agent associated with the determined context, wherein formulating the request in accordance with the grammar used by the domain agent includes:
>
>> *determining one or more required values and one or more optional values* associated with formulating the request in the grammar used by the domain agent;
>>
>> *extracting one or more criteria and one or more parameters* from one or more keywords contained in the one or more recognized words or phrases, wherein the parser extracts the one or more criteria and the one or more parameters using

6

procedures sensitive to the determined context;

inferring one or more further criteria and one or more further parameters associated with the request using a dynamic set of prior probabilities or fuzzy possibilities; and

transforming the one or more extracted criteria, the one or more extracted parameters, the one or more inferred criteria, and the one or more inferred parameters *into one or more tokens having a format compatible with the grammar used by the domain agent*, wherein the one or more tokens include all the required values and one or more of the optional values associated with formulating the request in the grammar used by the domain agent;

processing the formulated request with the domain agent associated with the determined context to generate a response to the utterance; and

presenting the generated response to the utterance via the speech unit.

The claim, albeit long, lacks any concrete or technological components to achieve the recited goal

of receiving and responding to a spoken request and is aspirational.

Finally, claim 1 of the '957 patent recites generic "modules" that receive a spoken request,

determine the words of the request, and identify the context of the request. ('957 patent at claim

1.) The "modules" also purportedly build a library of contexts corresponding to prior requests

(which the patent calls a "context stack") and compare the prior contexts to the words in the current

request:

1. A system for processing a natural language utterance, the system including one or more processors executing one or more computer program modules which, when executed, cause the one or more processors to:

generate a context stack comprising context information that corresponds to a plurality of prior utterances, wherein the context stack includes a plurality of context entries;

receive the natural language utterance, wherein the natural language utterance is associated with a command or is associated with a request;

determine one or more words of the natural language utterance by performing speech recognition on the natural language utterance;

identify, from among the plurality of context entries, *one or more context entries that correspond to the one or more words*, wherein the context information includes the one or more context entries, wherein identifying the one or more context entries comprises:

comparing the plurality of context entries to the one or more words;

generating, based on the comparison, *one or more rank scores* for individual

7

> context entries of the plurality of context entries; and
>
> *identifying, based on the one or more rank scores, the one or more context entries* from among the plurality of context entries; and

determine, based on the determined one or more words and the context information, *the command or the request* associated with the natural language utterance.

In other words, the claim recites receiving a request, somehow determining its words, somehow identifying its context by performing a comparison with contexts associated with prior requests, and somehow determining the command—all without any details of a particular technology or implementation.

Not only do the claims of the Context Patents lack any specific technological solution but so do their specifications. Indeed, the specifications recognize that speech recognition technology existed before the patents and was "successfully used in a wide range of applications." ('845 patent at 1:60-2:4.) They describe off-the-shelf "microphones (or other devices)" that receive a voice input. (*Id.* at

Figure 5. Speech Processing System Block Diagram
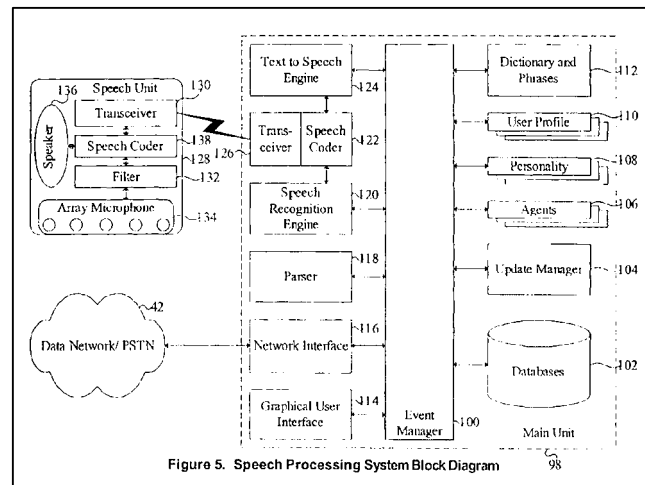
8:60-62; *see also* '006 patent at 7:33-35 (describing use of "array microphones to receive human speech").) And they identify computers with generically-labeled software components or "modules" (shown in Figure 5 of the '845 patent, at right) as performing functions on this input, including a "speech recognition engine" or "module," a "parser for parsing," a "network interface," and a "dictionary and phrases module." ('845 patent at 3:18-41, Fig. 5; *see also* '957 patent at 3:18-25, Fig. 5.) The components and modules are purely functional, and the specifications describe no technological details of any of them. ('845 patent at 3:42-51 (stating that system "takes action consistent with the command and performs other functions as detailed herein"), Fig. 5; *see* '957

8

patent at 3:30-35 ("this software may be distributed in any way . . . without altering the functions, features, scope, or intent of the invention"), Fig. 5.)

For example, the "domain agents" are described only functionally: they "include the behavior and data required for a specific area of functionality," *i.e.*, they respond somehow to requests that fall within a specific domain or field of inquiry. (*Id.* at 4:43-46; *see* '957 patent at 17:12-15, 27:66-28:4.) The specifications describe them as performing any imaginable functionality, from fast food ordering, to controlling windshield wipers on a car, to locating and playing music, with no details on *how* they perform any of these tasks. ('845 patent at 4:48-5:2; *see* '957 patent at 17:24-34.) The "parser" determines a "context for an utterance" using "prior probabilities or fuzzy possibilities." ('845 patent at 27:60-65.) But the patents do not explain how. Instead, they recognize that "probabilistic and fuzzy set decision and matching methods"—about which the specifications provide no details—are an (unspecified) way of determining what type of response the user desires based on the domain, context, and prior interactions with the user, as one would do in everyday conversation. (*Id.* at 7:9-23.) The "grammar" of the '006 patent claim is merely a format for packaging requests so that they can be processed by the corresponding domain agent, about which the specifications provide no details. ('006 patent at 18:12-16.) The "tokens" of the '006 patent are any pieces of information that are used to generate a response to a request. (*Id.* at 19:62-20:3.) And as for the "rank score" limitations of the '957 patent claim, the specifications do not even use that term. The '957 patent specification states that context entries may be compared against words in the request, and then "agents" associated with the context entries may be "ranked based on the determined score." ('957 patent at 20:31-46; *see also id.* at 35:7-15.) But it discloses no solution for this process, nor does it suggest that using computers for rank scoring— *i.e.*, comparing information—is anything but conventional computer processing.

9

The Context Patents do not describe any specific way to accomplish the goals of recogniz-

ing speech, determining a context, or generating a response as recited in the claims. Indeed, the

patents are agnostic as to any specific technological implementation, as confirmed by the '845 and

'720 patent specification: "*[r]egardless of the method of deployment[,]* the invention provides the

required functionality." ('845 patent at 20:2-3 (emphasis added).)

B.   The Multimodal Patents

The Multimodal Patents (the '039 and '468 patents) share much of their specifications and

have the same inventors.[5] These patents are directed to the same speech recognition, interpretation,

and response functions as the Context Patents, but performed on a multimodal request that has

both speech and non-speech portions. ('039 patent at claim 13; '468 patent at claim 19.)

Dialect asserts claim 13 of the '039 patent and claim 19 of the '468 patent in its complaint.

(Am. Compl. ¶¶ 213, 290.) Claim 13 of the '039 patent recites functional method steps: (1) re-

ceiving speech and non-speech communications; (2) transcribing each communication into text,

merging the texts, and somehow searching for text combinations; (3) determining a domain agent;

and (4) generating a response. To determine the domain agent, the claim recites using an unspec-

ified "context description grammar" that contains unspecified "entries." ('039 patent at claim 13.)

These entries are somehow given a relevance score after they are compared to text combinations

found in the transcribed and merged request, and that score is somehow used to select a domain

agent. (*Id.*) Claim 13 of the '039 patent states:

> 13. A method of processing speech and non-speech communications, comprising:
> *receiving the speech and non-speech communications*;
> *transcribing the speech and non-speech communications* to create a speech-based textual message and a non-speech-based textual message;
> *merging the speech-based textual message and the non-speech-based textual message* to

---

[5]   All cited disclosures in the '039 patent also appear in the '468 patent.

generate a query;

searching the query for text combinations;

comparing the text combinations to entries in a context description grammar;

accessing a plurality of domain agents that are associated with the context description grammar;

generating a relevance score based on results from comparing the text combinations to entries in the context description grammar;

selecting one or more domain agents based on results from the relevance score;

obtaining content that is gathered by the selected domain agents; and

generating a response from the content, wherein the content is arranged in a selected order based on results from the relevance score.

Claim 19 of the '468 patent recites a "method for processing multi-modal natural language inputs" with similar functional steps: (1) receiving a speech input and a non-speech input from a user; (2) transcribing the inputs and merging them; (3) determining a most likely context and a domain agent for that context; and (4) generating a response based on content provided by the domain agent. ('468 patent at claim 19.) Like the "entries in a context description grammar" of the '039 patent, the '468 patent recites using a "context stack" containing unspecified "entries" that are compared against the merged transcription to determine a most likely context. (*Id.*) And the '468 patent adds the idea of somehow identifying the user and using an unspecified "semantic knowledge-based model" with information about (a) prior interactions with the user, (b) prior interactions with other users, and (c) the user's environment, to transcribe the speech input. (*Id.*) The claim recites only black boxes to perform these steps: a "transcription module" that transcribes the non-speech input and a "speech recognition engine" that transcribes the speech input. (*Id.*) Claim 19 states:

19. A method for processing multi-modal natural language inputs, comprising:

receiving a multi-modal natural language input at a conversational voice user interface, the multi-modal input *including a natural language utterance and a non-speech input* provided by a user, wherein a transcription module coupled to the conversational voice user interface transcribes the non-speech input to create a non-speech-based transcription;

11

identifying the user that provided the multi-modal input;

creating a speech-based transcription of the natural language utterance using a speech recognition engine and a semantic knowledge-based model, wherein the semantic knowledge-based model includes a personalized cognitive model derived from one or more prior interactions between the identified user and the conversational voice user interface, a general cognitive model derived from one or more prior interactions between a plurality of users and the conversational voice user interface, and an environmental model derived from an environment of the identified user and the conversational voice user interface;

merging the speech-based transcription and the non-speech-based transcription to create a merged transcription;

identifying one or more entries in a context stack matching information contained in the merged transcription;

determining a most likely context for the multi-modal input based on the identified entries;

identifying a domain agent associated with the most likely context for the multi-modal input;

communicating a request to the identified domain agent; and

generating a response to the user from content provided by the identified domain agent as a result of processing the request.

While claim 13 of the '039 patent and claim 19 of the '468 patent may appear detailed at first glance, they are not. They boil down to the basic steps of receiving a request with speech and non-speech portions, transcribing it into text, somehow determining a context or domain, and responding to the request. Neither recites a specific technological solution.

The specifications of the Multimodal Patents recite no specific technological solution to perform the functions of the claims either. Instead, they describe receiving requests through admittedly "conventional" interfaces such as touch screens and keypads. ('468 patent at 1:28-36, 16:15-21; *see* '039 patent at 12:3-7, 18:63-19:14 (receiving spoken utterances at a "microphone array" and generic "analog or digital . . . speech interface").) The specifications functionally describe a "multi-pass automatic speech recognition module" that transcribes a speech input to text using whatever technique "is . . . available," such as preexisting "dictation grammar," "large vocabulary grammar," or "virtual dictation grammar" techniques. ('039 patent at 13:46-60; *see* '468 patent at 23:56-61 ("transcription may be performed using one or more conversational grammars,

12

a context description grammar 112," or any "other transcription methods").) The specifications only superficially describe the "virtual dictation grammar" as having "decoy words for out-of-vocabulary words" without explaining how it transcribes speech to text, and they disclose nothing about how to implement the other "grammars" mentioned in name only. ('039 patent at 13:52-60; *see* '468 patent at 23:56-61.) Nowhere do they suggest that the transcription techniques are anything but conventional.

The transcribing step of the '468 patent claim recites a "semantic knowledge-based model" that the specification describes as three black boxes: a "personalized cognitive model," a "general cognitive model," and an "environmental model" (shown in Fig. 8 of the '468 patent, at right). The patent does not explain how these black boxes transcribe a spoken input into text as the claim requires. (*See* '468 patent at claim 19.) Instead, it states that any information about prior interactions with the user, general prior interactions with other users, or the user's environment,

FIGURE 8

somehow "may be used to predict a user's next action" or "to enhance the interpretation of questions and/or commands," as humans have always done in conversations. (*Id.* at 23:64-24:48.)
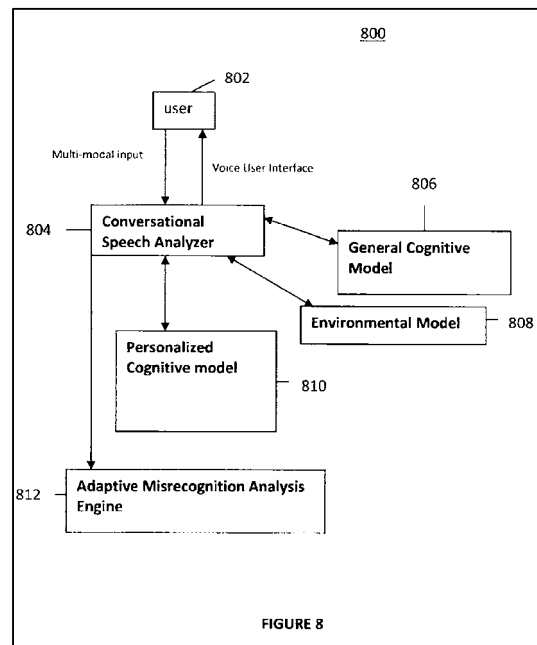
The specifications recite other generic software modules, including a "context description grammar module 112" (*id.* at 13:64-14:1, Fig. 1), a "context stack" (*id.* at 14:16-19), and a "response generator module" (*id.* at 3:24-30). These software modules are described only by their function. (*Id.* at 13:64-14:1, 14:6-7, 14:16-21 (describing functions of "context description grammar," "context stack," "response generator module").) The entries in the "context description

13

grammar" of the '039 patent claim are unspecified formats for how certain requests can be phrased; a request "may be compared against" these formats to determine, somehow, what is being asked for. ('039 patent at 13:64-14:1.) The specification of the '039 patent does not mention the "relevance score" recited in its claim. It merely states: "If a match is identified between an active grammar in the context description grammar and the command and/or request, then the match may be scored. The agents 106 may be ranked based on the determined score." (*Id.* at 13:66-14:2.) The entries in the "context stack" of the '468 patent claim are similar to the entries in the "context description grammar"; they are collections of "questions," "keywords and criteria" that can be matched to a request to somehow determine a most likely context. ('468 patent at 32:55-64, 20:19-28 (request may be compared to "a set of expected contexts that are stored in a context stack to determine a most likely context").) The specification discloses no specific solution for *how* to implement this matching to yield a "most likely context." And, again, the "domain agents" are unspecified software "packages or modules" containing unspecified "data that is needed to extend or modify the functionality of the system" to respond to requests in a specific domain or field of inquiry. ('039 patent at 14:38-47.)

III.   ARGUMENT

The Supreme Court's opinion in *Alice* directs courts to take a two-step approach. *Alice Corp. Pty. Ltd. v. CLS Bank Int'l*, 573 U.S. 208, 217-18 (2014). At Step One, a court must ask whether the claims are directed to a patent-ineligible abstract idea. *Id.* If they are, the court must then decide at Step Two whether the claims add an "inventive concept"—"an element or combination of elements that is 'sufficient to ensure that the patent in practice amounts to significantly more than a patent upon the [abstract idea] itself.'" *Id.* at 217-18 (quoting *Mayo Collaborative Servs. v. Prometheus Labs.*, 566 U.S. 66, 72-73 (2012). Unless these additional elements add something significant to the abstract idea, the claims are ineligible and invalid. *Id.* The Federal

Circuit has "repeatedly recognized that in many cases it is possible and proper to determine patent eligibility under 35 U.S.C. § 101 on a Rule 12(b)(6) motion." *Genetic Techs. Ltd. v. Merial L.L.C.*, 818 F.3d 1369, 1373 (Fed. Cir. 2016).

Dialect's patents are directed to variations on the abstract idea of responding to spoken requests. None recites a technological solution for achieving the claimed result as required to survive a Section 101 challenge. *In re TLI Commc'ns LLC Patent Litig.*, 823 F.3d 607, 611-12 (Fed. Cir. 2016) (affirming this Court's holding that claims directed to classifying and storing digital images were invalid under Section 101); *see also Two-Way Media Ltd. v. Comcast Cable Commc'ns, LLC*, 874 F.3d 1329, 1337 (Fed. Cir. 2017) (claims that are "directed to a result or effect that itself is the abstract idea and merely invoke[] generic processes and machinery" are ineligible under Section 101). Nor can Dialect save the claims through its amended pleading, which adds only conclusory allegations that are irrelevant, contrary to the patents themselves, or both. The Court should find the patents invalid.

> **A.**   **The Context Patents Are Directed to the Abstract Idea of Understanding and Responding to a Spoken Request Using Context and Recite No Inventive Concept.[6]**

In assessing whether patent claims are directed to an abstract idea at *Alice* Step One, the Court examines the "focus of the claims" or their "character as a whole." *BSG Tech LLC v. BuySeasons, Inc.*, 899 F.3d 1281, 1286 (Fed. Cir. 2018); *Elec. Power Grp., LLC v. Alstom S.A.*, 830 F.3d 1350, 1353 (Fed. Cir. 2016).

The asserted claims of the first category of the Context Patents (the '845, '720, and '006

---

[6]   Dialect's amended complaint still alleges infringement of only one claim from each asserted patent. Thus, the Court's ineligibility judgment would extend to those asserted claims. *See Hantz Software, LLC v. Sage Intacct, Inc.*, No. 2022-1390, 2023 WL 2569956 (Fed. Cir. Mar. 20, 2023). But even if Dialect were to attempt to assert other independent or dependent claims—and it has not done so despite amending its complaint in response to Amazon's initial motion to dismiss— those claims would fail for the same reasons discussed herein.

patents) are directed to the abstract idea of responding a spoken request using context, while the asserted claim of the '957 patent is directed to understanding (but not necessarily responding to) a spoken request using context. These practices have long been a feature of human interaction, as the patents confirm. Indeed, the patents explain that a person asking a question "rel[ies] heavily on context and the domain knowledge of the person answering." ('845 patent at 1:50-53; *see also* '957 patent at 1:56-60.) The first category of the Context Patents claim nothing more than the abstract idea, claiming only aspirational functions such as "perform speech recognition," "interpret the speech" or "recognize[] . . . words or phrases," "formulate a command," "create a response," and "execute the command" with no guidance about *how* to achieve those results. ('845 patent at claim 1; '720 patent at claim 1; '006 patent at claim 5); *BSG Tech*, 899 F.3d at 1285 ("If a claimed invention only performs an abstract idea on a generic computer, the invention is directed to an abstract idea at step one" of *Alice*). And claim 1 of the '957 patent recites result-oriented computer functions without any specific technology to achieve them: "computer program modules" that "generate a context stack" with "context entries," "receive" and interpret a natural language utter-ance, identify "context entries that correspond to" the recognized words, and "determine" the com-mand. ('957 patent at claim 1.) These result-oriented claims are unpatentable. *Affinity Labs of Tex., LLC v. Amazon.com Inc.*, 838 F.3d 1266, 1269 (Fed. Cir. 2016); *Apple, Inc. v. Ameranth, Inc.*, 842 F.3d 1229, 1244-45 (Fed. Cir. 2016); *IPA Techs., Inc. v. Amazon.com, Inc.*, 307 F. Supp. 3d 356, 363, 368-69 (D. Del. 2018).

At *Alice* Step One, the claims must "sufficiently describe how to achieve [the solution] in a non-abstract way." *Two-Way Media*, 874 F.3d at 1337. These claims do not. They recite only generic computing components configured or programmed to do what a *human* does: recognize speech, use context to understand what the speaker means, and—for the '845, '720, and '006

16

patents—respond. ('845 patent at 2:58-61 ("The environment makes significant use of context, prior information, domain knowledge, and user specific profile data to achieve a natural environment for one or more users making queries or commands in multiple domains."); *see also* '957 patent at 23:59-63.) Claim 1 of the '845 patent, for example, recites a computer processor that is "programmed to" perform various functions, including receiving a spoken request, performing speech recognition, interpreting the recognized speech, determining the context, formulating a command based on the context, determining where the command is to be executed, and executing it. Claim 1 of the '720 patent takes these functions and proclaims that they are performed by purely functional modules such as a "speech unit" that receives speech, a "speech recognition engine" that recognizes it, a "parser" that interprets it using "domain agents" chosen based on context, and an "agent architecture" that purportedly allows the domain agents to respond with no details on *how* any of these modules accomplish their functions. Claim 5 of the '006 patent does not even attempt to dress up the abstract idea. It recites a method of performing the steps of speech recognition, understanding, and response, without any technical detail that would save the claim from abstractness. And claim 1 of the '957 patent omits the final step of generating a response, while purporting to add a "rank[ing]" step where the current request is compared to contexts associated with prior requests—again without any details about *how* to implement this aspirational goal as a technical matter. Proclaiming that a computer should act like a human—*i.e.*, understand and respond to speech using "context"—is not an invention. It is a mere restatement of the difficult problem to be solved. *Epic IP LLC v. Backblaze, Inc.*, 351 F. Supp. 3d 733, 740 (D. Del. 2018) (invalidating claims directed to "individuals getting together to communicate"); *Gabara v. Facebook, Inc.*, 484 F. Supp. 3d 118, 127-28 (S.D.N.Y. 2020) (invalidating claims directed to "identifying topics in a conversation, searching the internet for those topics," and generating new topics

17

from the search results), *aff'd*, 852 F. App'x 541 (Fed. Cir. 2021).

Courts have repeatedly invalidated similar patents that fail to claim a specific technological solution. For example, the claims in *Apple*, 842 F.3d at 1244-45, were directed to using voice capture to generate menus. The Federal Circuit held that merely reciting "voice capture technologies" did not disclose "how these elements were to be technologically implemented." *Id.* at 1245. The claim was thus abstract. In *IPA*, the court rejected claims directed to the abstract idea of transmitting electronic data to a user in response to a spoken request. *IPA*, 307 F. Supp. 3d at 363. The claims in *IPA* recited "receiving," "interpreting," "translating" a spoken request to "search" a database, and "transmitting" results to the user, but were "aspirational in nature and devoid of any implementation details or technical description." *Id.* at 368-369. In *West View Research*, the Federal Circuit held that claims directed to a "computerized apparatus capable of interactive information exchange with a human user," including via "a microphone," were abstract. *W. View Rsch., LLC v. Audi AG*, 685 F. App'x 923, 925-26 (Fed. Cir. 2017). The claims recited "receiving" a data query, "analyzing" it, "retrieving and processing" information constituting a response, and "generating" a response—a "familiar class of claims 'directed to' a patent-ineligible concept." *Id.* at 926 (quoting *Elec. Power Grp.*, 830 F.3d at 1353). And just weeks ago, in *Trinity Info Media*, the Federal Circuit held that claims directed to "(1) receiving user information; (2) providing a polling question; (3) receiving and storing an answer; (4) comparing that answer to generate a 'likelihood of match' with other users; and (5) displaying certain user profiles based on that likelihood" were abstract at *Alice* Step One. *Trinity Info Media, LLC v. Covalent, Inc.*, 72 F.4th 1355, 1362 (Fed. Cir. 2023). The asserted claims of the Context Patents are similarly aspirational in nature, lack any implementation details, and are invalid. *See Enco Sys., Inc. v. DaVincia, LLC*, 845 F. App'x 953, 957 (Fed. Cir. 2021) (invalidating speech-to-text processing patent directed to

18

the abstract idea of automating the audio-video captioning process performed by humans).

Dialect cannot rely on the patents' specifications to save the asserted claims. That is because "the *claim*—as opposed to something purportedly described in the specification"—must disclose a specific technological solution. *Two-Way Media*, 874 F.3d at 1338. In any event, the shared portions of the Context Patents' specifications fail to even suggest the use of any specific, unconventional technology to carry out the claimed functions. Indeed, the specifications describe the elements of the alleged invention as software "modules" or "engines" that are programmed— somehow, some way—to carry out the aspirational idea of responding to a spoken request using context. (*E.g.*, '845 patent at 1:60-2:4, 3:18-51, 4:43-5:2 (describing "domain agents"), 15:63-16:55, 27:60-65 (describing "parser"); '006 patent at 18:12-16 (describing "grammar" of domain agent), 19:62-20:3 (describing "tokens" used to generate response to a request).) Indeed, the programming is left for others to devise. The specification of the '845 and '720 patents expressly disclaims any specific implementation of the functions recited in the claims. ('845 patent at 20:2-3 ("Regardless of the method of deployment[,] the invention provides the required functionality.").) And the specification of the '957 patent states only that "scoring" can be done to find a matching context ('957 patent at 4:47-51, 20:36-39, 33:40-44, 35:1-12, 36:44-47, 37:3-15) in "real-time" to allow "rapid responses" and a "natural question and response environment" (*id.* at 7:31-37) and, if the system cannot find an adequate match, it may ask further questions of the user (*id.* at 7:52-55, 35:12-15). Just as the specifications admit, humans perform each of these steps, many times over, in any conversation, every day. Thus, the Context Patents are directed to an abstract idea at *Alice* Step One.

The asserted claims of these patents also fail Step Two of the *Alice* test, as they recite no inventive concept. Generic computer technology, "well-understood, routine, conventional"

19

elements, or "purely functional" elements cannot supply the required inventive concept. *Alice*, 573 U.S. at 221-26 (citing *Mayo*, 566 U.S. at 72-73, 77, 82). Yet here, that is all the patents disclose. Neither the claims nor the specifications recite any new computer hardware. They also recite no new software, and in fact confirm that speech recognition on computers existed long before the patents. ('845 patent at 1:60-2:4.) As described above, the asserted claims of the '845, '720, and '006 patents recite generic processors and computing components performing the functions of receiving a request, interpreting it using context, and responding. ('845 patent at claim 1; '720 patent at claim 1; '006 patent at claim 5.) Claim 1 of the '957 patent recites similar steps of receiving and interpreting a request, using the idea of "scoring" to find a matching context. ('957 patent at claim 1.) The asserted claims of the Context Patents are silent about *how* any of these functions are performed or implemented. The absence of "how to" dooms the claims. *Affinity*, 838 F.3d at 1268 (claims that merely "stat[e] [the relevant] functions in general terms" lack inventive concept at *Alice* Step Two); *see also Hawk Tech. Sys., LLC v. Castle Retail, LLC*, 60 F.4th 1349, 1358 (Fed. Cir. 2023) (invalidating claims "recited at such a level of result-oriented generality that those claims amount to a mere implementation of an abstract idea"); *Dropbox, Inc. v. Synchronoss Techs., Inc.*, 815 F. App'x 529, 537 (Fed. Cir. 2020) (no inventive concept where the claims "provide[] less of 'a specific means or method' than 'a result or effect that itself is the abstract idea and merely invokes generic processes and machinery'"). In fact, the only difference between these claims and the abstract idea itself—responding to spoken requests using context— is that the claims are carried out on a computer. This is insufficient to save the claims at *Alice* Step Two. *Alice*, 573 U.S. at 222-23 (a "computer implementation" of an abstract idea does "not supply the necessary inventive concept"); *see also Brunswick Corp. v. Volvo Penta of the Ams., LLC*, No. 1:22-CV-00108, 2022 WL 16856094, at *10 (E.D. Va. Nov. 10, 2022) (claims purporting to

20

implement abstract idea on generic computers lack inventive concept).

Nor can Dialect avoid the conventional nature of these claims by asserting that they include a specific ordered combination of known elements. An ordered combination lacks an inventive concept where the "recited physical components behave exactly as expected according to their ordinary use" or are "organized in a completely conventional way." *TLI Commc'ns*, 823 F.3d at 615; *Two-Way Media*, 874 F.3d at 1341. For the '845, '720, and '006 patents, a request must be received before it can be interpreted, and it must be interpreted in the right context before a response can be provided. Likewise, for the '957 patent, a request must be received, the words must be recognized, and the words must be compared to contexts associated with prior requests before the request can be understood. These are the same fundamental processes that humans perform in any conversation. The stated goal of the patents is to take these fundamental processes and implement them on generic computers, without any specific disclosure of *how* one would do so as a technological matter. ('845 patent at 1:50-53; '957 patent at 1:56-60.) Aspirations are not patentable. The claims fail both steps of the *Alice* test and are invalid.

B. The Multimodal Patents Are Directed to the Abstract Idea of Responding to a Request that Includes Spoken and Non-Spoken Portions and Recite No Inventive Concept.

At *Alice* Step One, the asserted claims of the Multimodal Patents are directed to the same abstract idea of the Context Patents—responding to a request using context—with a slight variation. Instead of just a spoken request, the request contains both a spoken portion and a non-spoken portion. ('039 patent at claim 13 ("method of processing speech and non-speech communications"); '468 patent at claim 19 ("method for processing multi-modal natural language inputs").) The idea of these patents is no different than a human receiving spoken and non-spoken information—such as a student listening to a professor's lecture while viewing notes on a blackboard—and merging the two pieces of information in their mind or writing them down before responding.

21

Like the Context Patents, the Multimodal Patents aspirationally seek to have computers provide human-like, "natural responses" to natural language requests. ('039 patent at 1:39-40, 1:55-2:7.) The patents claim the result of computers responding to such requests by doing what a human does in everyday multimodal conversations: relying on context. (*Id.* at 2:12-21 (the invention "makes maximum use of context, prior information, domain knowledge, and user specific profile data to achieve a natural environment"), 1:40-44 (explaining that a person asking a question "relies heavily on context and domain knowledge" of the person they are speaking to).)

The asserted claims do not recite any specific technology for using context to respond to a multimodal request, and thus they fail to describe how to achieve this aspirational result in a non-abstract way. *See Two-Way Media*, 874 F.3d at 1337. Asserted claim 13 of the '039 patent recites only result-oriented functions that have long been performed by humans: receiving spoken and non-spoken portions of a request, merging them into a single message, comparing the message to known request formats (entries in the "context description grammar"), identifying the field of inquiry and desired response type (the "domain agent") based on that comparison, and responding. ('039 patent at claim 13.) Asserted claim 19 of the '468 patent recites similar human functions as being performed by generic black boxes—for example, a "transcription module" and "speech recognition engine." ('468 patent at claim 19.) Such result-oriented claims are unpatentable. *Trinity Info Media*, 72 F.4th at 1362; *Affinity*, 838 F.3d at 1269; *Apple*, 842 F.3d at 1244-45; *IPA*, 307 F. Supp. 3d at 363, 368-69 (claims reciting "receiving," "interpreting," and "translating" a spoken request were "aspirational in nature and devoid of [] implementation details or [] description.").

Receiving a request that has spoken and non-spoken portions does not change the abstract nature of the claims. The asserted claim 13 of the '039 patent does not recite any specific way of

"receiving the speech and non-speech communications." The asserted claim 19 of the '468 patent

recites only a black-box "conversational voice user interface" to receive the request. The patents

confirm that speech and non-speech interfaces were conventional technology, and prior art

acknowledged in the patents themselves confirms that multimodal conversation was known at the

time. ('039 patent at Other Publications; '468 patent at Other Publications, 1:28-36, 16:15-21.)

Nor does the "semantic knowledge-based model" claimed in the '468 patent provide any

specific technological solution for transcribing speech to text. The claim simply states that the

"model" includes unspecified information about prior interactions with the user, prior interactions

with other users, and the environment—all broad categories of information that humans use in

everyday conversation—with no solution for how to use any of them. ('468 patent at claim 19;

see id. at 2:8-11.)

The specifications of the Multimodal Patents cannot cure the abstract nature of the claims.

Two-Way Media, 874 F.3d at 1338. But, in any event, the specifications also disclose no specific

solution for achieving the claimed result of responding to a multimodal request using context.

They identify only preexisting technology and recite generically-labeled software "modules" that

carry out the aspirational functions recited in the claims in an undisclosed way. ('039 patent at

13:46-60 (reciting a "speech recognition module" that uses preexisting speech-to-text transcription

techniques), 13:64-14:1 (reciting a "context description grammar" without any description of what

it is or how it is programmed), 14:16-23 (reciting a "context stack" with no description of how it

is created), 3:24-30 (reciting a "response generator module" that generates a response, without any

implementation details); '468 patent at 1:28-36, 16:15-21 (reciting existing speech and non-speech

interfaces), 23:56-61 (reciting "speech recognition engines that transcribe utterances to textual

messages" using preexisting transcription methods), 24:40-48 (reciting use of "cognitive model"

23

and "environmental model" to "enhance responses to commands," with no description of how this is achieved).) Moreover, they repeatedly disclaim any specific technological implementation of these functions. ('468 patent at 3:23-27 (software modules "may be distributed in any way"), 17:45-54 ("The invention supports any suitable configuration [of servers]."), 23:33-35 (user interface components "may be distributed in any suitable manner"). The asserted claims of the patents are directed to an abstract idea in Step One of the *Alice* test.

The Multimodal Patents also fail *Alice* Step Two. Neither the claims nor the specifications disclose any new hardware or software for achieving the aspirational result of responding to a multimodal request using context. (*See, e.g.*, '039 patent at 13:46-60 (speech recognition module uses preexisting transcription techniques); '468 patent at 1:28-36, 16:15-21 (user interface may include conventional touch screen display or keypad).) Indeed, the patents confirm that computer-based "speech recognition" was "successfully used in a wide range of applications" before the patents. ('039 patent at 1:52-55.) The asserted claims recite only black-box modules and purely functional method steps—receiving spoken and non-spoken portions of a request, transcribing them into text and merging the text, comparing the request to known request formats, identifying a context or domain based on that comparison, and responding to the request—without explaining how these functions are carried out. ('039 patent at claim 13; '468 patent at claim 19.) These claims have no inventive concept. *Affinity*, 838 F.3d at 1269 (invalidating claims "describ[ing] a desired function or outcome, without providing any limiting detail . . . to a particular solution"); *Apple*, 842 F.3d at 1245 (claims that "refer[red] to the use of . . . voice capture technologies without providing how these elements were to be technologically implemented" were non-inventive); *Dropbox*, 815 F. App'x at 537.

Nor do the asserted claims provide any inventive concept in an ordered combination of

24

functional steps. These claims recite functions in the conventional order in which they would need to be performed. The request must be received before it is transcribed. Spoken and non-spoken portions must be converted into the same medium (*e.g.*, text) before they are merged, and they must be merged before determining the context or domain of the full request. And the context or domain must be identified before responding to the request. *TLI Commc'ns*, 823 F.3d at 615; *Two-Way Media*, 874 F.3d at 1341 (steps "organized in a completely conventional way" are not inventive).

As with the Context Patents, these functional claims seek to monopolize all ways of achieving their stated result, without any technological solution of how to do so. *Affinity*, 838 F.3d at 1268; *Hawk Tech.*, 60 F.4th at 1358. The asserted claims of the Multimodal Patents thus fail Step Two of the *Alice* test and are invalid under Section 101.

C.   Dialect's Purported Factual Allegations in the Amended Complaint Cannot Save the Asserted Claims.

Rather than respond to Amazon's initial motion to dismiss, Dialect filed an amended complaint adding: characterizations of the patents or alleged inventions as new; allegations about the problems to be solved by the patents; and allegations about research and development efforts, business discussions with a company called VoiceBox, and industry praise. Dialect cannot save the asserted claims from invalidity under Section 101 by amending its pleading.

Irrelevant allegations regarding novelty or non-obviousness. Dialect's amended complaint repeatedly points to allegations about novelty or non-obviousness, as though these allegations preclude a finding of invalidity under Section 101. (Am. Compl. ¶¶ 20-27 (describing third-party VoiceBox as having created "groundbreaking" technology); *id.* ¶¶ 28-31 (describing alleged meetings where VoiceBox presented "award-winning" technology); *id.* ¶¶ 45, 53, 69, 79, 87, 95 (alleging that the inventors conceived of "novel software techniques"); *id.* ¶¶ 48, 57, 72, 82, 90

(allegations regarding what the patent examiner did not locate in the prior art).) They do not. The Federal Circuit has repeatedly confirmed that novelty and patent eligibility are "distinct" concepts, as "a claim for a *new* abstract idea is still an abstract idea." *Synopsys, Inc. v. Mentor Graphics Corp.*, 839 F.3d 1138, 1151 (Fed. Cir. 2016). Thus, "merely reciting an abstract idea by itself in a claim—even if the idea is novel and non-obvious—is not enough to save it from ineligibility." *Solutran, Inc. v. Elavon, Inc.*, 931 F.3d 1161, 1169 (Fed. Cir. 2019); *see also Affinity Labs*, 838 F.3d at 1270 n.3 ("[T]he eligibility finding does not turn on the lack of novelty of the claim; it turns on the fact that the claim is drawn to any embodiment of an abstract idea.").

These controlling principles render Dialect's complaint amendments irrelevant. Dialect's allegations about "groundbreaking" or "award-winning" technology, or new software techniques, at most relate to alleged novelty and thus do not make the claims patent-eligible. Dialect's quotation of the reasons a patent examiner allowed the patent claims to issue over the prior art (Am. Compl. ¶¶ 48, 57, 72, 82, 90) are likewise irrelevant to eligibility. Indeed, the Federal Circuit has instructed that "conclusory allegations that the prior art lacked elements of the asserted claims are insufficient to demonstrate an inventive concept." *Trinity Info Media*, 72 F.4th at 1366. And even patent examiners' decisions on patent-eligibility—that are missing in the prosecution histories of the asserted patents—receive no deference. *Novo Nordisk A/S v. Caraco Pharm. Labs., Ltd.*, 719 F.3d 1346, 1357 (Fed. Cir. 2013) ("The initial determinations by the PTO in determining to grant the application are entitled to no deference."). Dialect cannot articulate any inventive, *technical* contribution beyond employing existing computer technology toward the aspirational goal of receiving, processing, and responding to spoken or multimodal requests.

Conclusory and irrelevant allegations about the problem to be solved. Dialect pleads that the claims address a "uniquely challenging technological problem" of interpreting human

communications, that they "overc[ame] the technical limitations of existing speech recognition systems," and that they include "novel combinations and ordering of techniques and structures" not found in the prior art. (Am. Compl. ¶¶ 42-44, 45, 51-53, 47, 56, 68-69, 71, 75-79, 81, 85, 87, 89, 93-95; *see also id.* ¶¶ 46, 54-55, 70, 80, 86, 88, 96 (similar).) But the Court need not accept such "legal conclusions as true, even if couched as factual allegations." *Bridge & Post, Inc. v. Verizon Communs., Inc.*, 778 F. App'x 882, 894 (Fed. Cir. 2019); *Dropbox*, 815 F. App'x at 538 (allegations that "restate the claim elements and append a conclusory statement that 'nothing in the specification describes these concepts as well-understood, routine, or conventional'" do not create factual dispute). For the same reasons, Dialect cannot rely on allegations purporting to describe the "difficult problem" of "building natural-language speech-recognition interfaces for computers and vehicles." (Am. Compl. ¶¶ 42-44; *see also id.* ¶¶ 51-52, 68, 75-78, 85-86, 93-94.) That is because, as explained above, the asserted claims fail to recite any *specific*, *technological* solution for providing natural, human-like responses to requests. *Supra* Sections III.A, III.B.

Dialect's new allegations also fail for the independent reason that the purported problems and solutions in the amended complaint do not appear in the claims as required. *See Two-Way Media*, 874 F.3d at 1338 (holding that the *Alice* inquiry turns on what is "*claim[ed]*—as opposed to something purportedly described in the specification"). For example, for all six patents subject to this motion, Dialect alleges that the inventions are a technological improvement because they interpret incomplete queries based on domain information, user preferences, or scoring. (Am. Compl. ¶¶ 44-45, 51-53, 67-69, 78-79, 86-87, 93-94.) But the asserted claims recite no such solution; none of the claims recites any incomplete query at all. Dialect alleges that the improvements of the '006 patent include "determining an identity of a user based on voice characteristics" (*id.* ¶¶ 53-54), "satisfying a predetermined confidence level" (*id.* ¶ 54), and "updating user specific

27

vocabularies or dictionaries" (*id.*). None of these appears in the claims. For the '468 patent, Dialect's amendment simply restates the functional results that are recited in the claim, such as "creat[ing] and merg[ing] transcripts of the inputs using a speech recognition engine." (*Id.* ¶ 69.) But neither the claim nor Dialect's conclusory amendment identifies *how* this would be done. Dialect alleges that the '845 patent resolves problems such as "retrieval of both local and network online information," "speed of interaction," difficulties with "wireless communications" in vehicles, and the "inherent danger and distraction" of driving a vehicle. (*Id.* ¶¶ 75-77.) But the claim, as noted above, recites only the functions of receiving and recognizing speech, determining its context, and executing the command. *Supra* Section II.A. And although Dialect points to the '957 patent as relating to "a text message" input that is processed "using a multi-pass automatic speech recognition module," by searching the text message for characters or words, and by scoring and ranking the results of a comparison (Am. Compl. ¶ 96), asserted claim 1 recites no such thing. Indeed, the claim specifically calls for a "natural language utterance," not a text message, and recites no "multi-pass" automatic speech recognition contrary to what Dialect's complaint suggests. *Supra* Section II.A. Dialect's reliance on unclaimed features or conclusory characterizations of the claims is irrelevant. *Two-Way Media*, 874 F.3d at 1338.

Irrelevant allegations regarding research, business discussions, and industry praise. Dialect's amended complaint also points to alleged research efforts, business discussions, or industry praise. (Am. Compl. ¶¶ 20-27 (third-party VoiceBox's purported research activities and industry recognition); *id.* ¶¶ 28-31 (alleged meetings between Amazon and VoiceBox); *id.* ¶¶ 32-38 (alleged copying and poaching of VoiceBox employees).) These allegations are irrelevant to patent eligibility. Indeed, Dialect's allegations about business discussions between VoiceBox and Amazon, replete with many pages of alleged screenshots from a meeting (*id.* ¶¶ 28-38), lack any

discernable connection to the asserted patent claims or the legal standard for patent eligibility. The

claims themselves recite no specific technological solution for the aspirational goal of receiving,

processing, and responding to spoken or multimodal requests.

> **D.   Dialect's Counts Relating to the Context Patents and the Multimodal Patents
> Should Be Dismissed with Prejudice.**

The Court should dismiss Dialect's claims of alleged infringement of the Context Patents

and the Multimodal Patents (Counts I, II, and IV-VII of the amended complaint) with prejudice,

as any further amendment of the complaint with respect to these counts would be futile. Dialect

has already amended its complaint after reviewing Amazon's initial motion to dismiss. But as

discussed above, none of the amendments can save the claims from a finding of invalidity.

The patent claims themselves and the accompanying disclosures of the specifications allow

the Court to resolve the patent-eligibility question as a matter of law. *Simio, LLC v. FlexSim*

Software Prod., Inc., 983 F.3d 1353, 1365 (Fed. Cir. 2020); *WhitServe LLC v. Donuts Inc.*, 809 F.

App'x 929, 935 (Fed. Cir. 2020) (patent eligibility may be determined at the pleading stage where,

for example, the specification confirms that the claims are abstract and lack inventive concept).

No discovery is needed to resolve this motion, as the patents fail to disclose any new computer

hardware and admit that they seek to carry out or replicate natural human activity on preexisting

computers. *Cleveland Clinic Found. v. True Health Diagnostics LLC*, 859 F.3d 1352, 1360 (Fed.

Cir. 2017) ("[W]e have repeatedly affirmed § 101 rejections at the motion to dismiss stage, before

claim construction or significant discovery has commenced."); *see also Trinity Info Media*, 72

F.4th at 1360-61 (rejecting argument that subject matter eligibility challenge could not proceed

without discovery or claim construction, and finding claims ineligible under § 101). Nor is any

discovery needed to understand that the order of steps claimed in the patents—*i.e.*, receiving a

request, interpreting it, and responding to it—is no different from what humans do, day in and day

out. Dialect has taken its chance to amend the complaint; no further amendment could change the

abstract and ineligible nature of the claims. *WhitServe*, 809 F. App'x at 935; *see also Simio*, 983

F.3d at 1364 (affirming "futility-based denial" of leave to amend).

IV.   CONCLUSION

Amazon requests that the Court dismiss Counts I, II, and IV-VII of the amended complaint

for failure to allege infringement of a patentable claim under 35 U.S.C. § 101.

Dated: August 14, 2023   Respectfully submitted,

                                                     /s/ Laura Anne Kuykendall
                                                     Robert A. Angle (VSB No. 37691)
                                                     robert.angle@troutman.com
                                                     Laura Anne Kuykendall (VSB No. 82318)
                                                     la.kuykendall@troutman.com
                                                     TROUTMAN PEPPER HAMILTON SANDERS
                                                     LLP
                                                     1001 Haxall Point
                                                     Richmond, Virginia 23219
                                                     Telephone: (804) 697-1468
                                                     Facsimile: (804) 697-1339

                                                     Mary C. Zinsner (VSB No. 31397)
                                                     mary.zinsner@troutman.com
                                                     TROUTMAN PEPPER HAMILTON SANDERS
                                                     LLP
                                                     401 9th Street NW, Suite 1000
                                                     Washington, DC 20004
                                                     Telephone: (202) 274-1932
                                                     Facsimile: (202) 274-2994

                                                     Of counsel:

                                                     J. David Hadden, *pro hac vice*
                                                     dhadden@fenwick.com
                                                     Saina S. Shamilov, *pro hac vice*
                                                     sshamilov@fenwick.com
                                                     Ravi R. Ranganath, *pro hac vice*
                                                     rranganath@fenwick.com
                                                     Vigen Salmastlian, *pro hac vice*
                                                     vsalmastlian@fenwick.com
                                                     Donna T. Long, *pro hac vice*

dlong@fenwick.com
Johnathan Lee Chai, *pro hac vice*
jchai@fenwick.com
FENWICK & WEST LLP
801 California Street
Mountain View, CA 94041
Telephone: (650) 988-8500
Facsimile: (650) 938-5200

Jonathan T. McMichael, *pro hac vice*
jmcmichael@fenwick.com
FENWICK & WEST LLP
401 Union Street, 5th Floor
Seattle, WA 98101
Telephone: (206) 389-4510
Facsimile: (206) 389-4511

Daniel W. Ledesma, *pro hac vice*
dledesma@fenwick.com
FENWICK & WEST LLP
902 Broadway, 18th Floor
New York, NY 10010-6035
Telephone: (212) 430-2600
Facsimile: (650) 938-5200

Counsel for Defendants Amazon.com, Inc., et al.

31